OCR
Oxford Cambridge and RSA

Monday 23 May 2022 – Afternoon

AS Level Further Mathematics A

Y532/01 Statistics

Time allowed: 1 hour 15 minutes

You must have:
- the Printed Answer Booklet
- the Formulae Booklet for AS Level Further Mathematics A
- a scientific or graphical calculator

INSTRUCTIONS
- Use black ink. You can use an HB pencil, but only for graphs and diagrams.
- Write your answer to each question in the space provided in the **Printed Answer Booklet**. If you need extra space use the lined pages at the end of the Printed Answer Booklet. The question numbers must be clearly shown.
- Fill in the boxes on the front of the Printed Answer Booklet.
- Answer **all** the questions.
- Where appropriate, your answer should be supported with working. Marks might be given for using a correct method, even if your answer is wrong.
- Give non-exact numerical answers correct to **3** significant figures unless a different degree of accuracy is specified in the question.
- The acceleration due to gravity is denoted by $g\,\mathrm{m\,s^{-2}}$. When a numerical value is needed use $g = 9.8$ unless a different value is specified in the question.
- Do **not** send this Question Paper for marking. Keep it in the centre or recycle it.

INFORMATION
- The total mark for this paper is **60**.
- The marks for each question are shown in brackets **[ ]**.
- This document has **8** pages.

ADVICE
- Read each question carefully before you start your answer.

Answer **all** the questions.

1 A geography student chose a certain point in a stream and took measurements of the speed of flow, $v\,\text{ms}^{-1}$, of water at various depths, $d\,\text{m}$, below the surface at that point. The results are shown in the table.

d	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
v	0.8	0.5	0.7	1.2	1.1	1.3	1.6	1.4	0.4

$n = 9 \qquad \Sigma d = 2.7 \qquad \Sigma v = 9.0 \qquad \Sigma d^2 = 0.96 \qquad \Sigma v^2 = 10.4 \qquad \Sigma dv = 2.85$

(a) (i) Explain why d is an example of an independent, controlled variable. **[1]**

(ii) Use **two** relevant terms to describe the variable v in a similar way. **[1]**

A statistician believes that the point (0.5, 0.4) may be an anomaly.

(b) Calculate the equation of the least squares regression line of v on d for all the points in the table **apart from** (0.5, 0.4). **[2]**

(c) Use the equation of the line found in part **(b)** to estimate the value of v when $d = 0.5$. **[1]**

(d) Use your answer to part **(c)** to comment on the statistician's belief. **[1]**

(e) Use the diagram in the Printed Answer Booklet (which does **not** illustrate the data in this question) to explain what is meant by "least squares regression line". **[2]**

2 Eight runners took part in two races. The positions in which the runners finished in the two races are shown in the table.

Runner	A	B	C	D	E	F	G	H
First race	3	1	5	6	2	8	7	4
Second race	4	3	8	7	2	5	6	1

Test at the 5% significance level whether those runners who do better in one race tend to do better in the other. **[7]**

3 A discrete random variable X has the following probability distribution.

x	1	2	3	4
$\text{P}(X = x)$	p	0.31	0.3	p^2

(a) Determine the value of p. **[3]**

(b) It is given that $\text{E}(aX + b) = \text{Var}(aX + b) = 23.19$, where a and b are positive constants. Determine the value of a and the value of b. **[6]**

4 A school pupil keeps a note of whether her journeys to school and from school are delayed. The results for a random sample of journeys are shown in the table.

	Direction of journey	
	To school	From school
Delayed	64	56
Not delayed	74	106

Test at the 10% significance level whether there is association between delays and the direction of the journey. **[7]**

5 The manager of an emergency response hotline believes that calls are made to the hotline independently and at constant average rate throughout the day. From a small random sample of the population, the manager finds that the mean number of calls made in a 1-hour period is 14.4.

Let R denote the number of calls made in a randomly chosen 1-hour period.

(a) Using evidence from the small sample, state a suitable distribution with which to model R. You should give the value(s) of any parameter(s). **[1]**

(b) In this part of the question, use the distribution and value(s) of the parameter(s) from your answer to part **(a)**.

(i) Find $\mathrm{P}(R > 20)$. **[2]**

(ii) Given that $\mathrm{P}(R = r) > \mathrm{P}(R = r+1)$, show algebraically that $r > 13.4$. **[2]**

(iii) Hence write down the mode of the distribution. **[1]**

The manager also finds, from records over many years, that the modal value of R is 10.

(c) Use this result to comment on the validity of the distribution used in part **(b)**. **[1]**

(d) Assume now that the type of distribution used in part **(b)** is valid. Find the range(s) of values of the parameter(s) of this distribution that would correspond to the modal value of R being 10. **[2]**

6 A teacher has 10 different mathematics books. Of these books, 5 are on Algebra, 3 are on Calculus and 2 are on Trigonometry.

The teacher chooses 5 of the books at random.

(a) Find the probability that 3 of the books are on Algebra. **[3]**

The teacher now arranges all 10 books in random order on a shelf.

(b) Find the probability that the Calculus books are next to each other and the Trigonometry books are next to each other. **[3]**

In this question you must show detailed reasoning.

(c) Find the probability that 2 of the Calculus books are next to each other but the third Calculus book is separated from the other 2 by at least 1 other book. **[4]**

7 Each of three students, X, Y and Z, was given an identical pack of 48 cards, of which 12 cards were red and 36 were blue. They were each told to carry out a different experiment, as follows:

Student X: Choose a card from the pack, at random, 20 times altogether, **with** replacement. Record how many times you obtain a red card.

Student Y: Choose a card from the pack, at random, 20 times altogether, **without** replacement. Record how many times you obtain a red card.

Student Z: Choose single cards from the pack at random, **with** replacement, until you obtain the first red card. Record how many cards you have chosen, including the first red card.

(a) Find the probability that student Z has to choose more than 8 cards in order to obtain the first red card. **[3]**

Each student carries out their experiment 30 times.

The frequencies of the results recorded by each student are shown in the following table, but **not** necessarily with the rows in the order X, Y, Z:

	Number recorded	0	1	2	3	4	5	6	7	8	$\geqslant 9$	Observed Mean	Observed Variance
Observed Frequencies	Student 1	0	0	1	3	7	8	6	4	1	0	5.03	1.97
	Student 2	0	8	5	4	2	3	3	2	1	2	4.03	11.57
	Student 3	0	1	2	5	4	6	5	3	4	0	4.97	3.70

(b) In this question you must show detailed reasoning.

Two other students make the following statements about the results. For each of the statements, explain whether you agree with the statement. Do **not** carry out any hypothesis tests, but in each case you should give **two** justifications for your answer.

(i) "The second row is a good match with the expected results for student Z." **[4]**

(ii) "The third row is definitely student X's results." **[3]**

END OF QUESTION PAPER

BLANK PAGE

BLANK PAGE